

Sensory neuroscience: Visualizing the auditory cortex

Andrew J. King and Jan W.H. Schnupp

Recent studies suggest that the auditory cortex may use sensory processing strategies analogous to those already established for the visual cortex. Nevertheless, fundamental differences in the way the visual and auditory worlds are structured have to be borne in mind.

Address: University Laboratory of Physiology, Parks Road, Oxford OX1 3PT, UK.

Current Biology 1998, 8:R784–R787
<http://biomednet.com/elecref/09609822008R0784>

© Current Biology Ltd ISSN 0960-9822

Characterizing the stimulus selectivity of sensory neurons is a prerequisite for understanding how the world is represented within the brain. It is also an essential step in identifying the contributions of different brain areas to particular perceptual tasks. The most extensively studied sensory neurons are those of the primary visual cortex (V1). Because the surface of the retina is mapped onto that of the cortex, V1 contains an orderly representation of visual space and each neuron responds to visual stimuli within a restricted area of space — its ‘classical receptive field’. Hubel and Wiesel [1] demonstrated nearly 40 years ago that V1 neurons do not respond to diffuse illumination within their receptive fields, but are instead selective for lines or bars of a particular orientation. The preferred orientation remains constant within columns of neurons that span the layers of the cortex, but varies smoothly from one column to the next across the cortical surface. Consequently, visual scenes are mapped onto the cortex as an array of short, oriented lines.

Neurons in V1 also vary in their preferences for other stimulus attributes. These include the relative extent to which a stimulus activates the two eyes (‘ocular dominance’) and the disparity between the image formed on each eye — the basis of stereoscopic vision — the spatial frequencies contained in the stimulus and its direction of motion. Global measurements of cortical activity, provided by optical imaging techniques, suggest that the representations of different stimulus features are superimposed and interwoven in a complex manner within the map of visual space [2].

For years, some researchers have suspected that the functional organization of the primary auditory cortex (A1) may be similar to that of V1. But progress in this area has been slow, not least because many of the simple auditory stimuli used to study A1, such as pure tones, noises and clicks, are not particularly effective in activating these neurons. In general, researchers have lacked stimuli that

are ‘optimal’ for the study of A1, in the sense that moving, oriented bars appear to be optimal for V1. Nevertheless, some parallels between A1 and V1 can be drawn. Again we find that the peripheral receptor surface is mapped over the surface of the cortex. But, unlike the eye, the inner ear maps sound frequency, not space, along its receptor surface. This ‘tonotopic’ order is preserved in the projections from the cochlea to A1, and the surface of A1 can be thought of as a series of ‘iso-frequency bands’.

The preferred sound frequency of A1 neurons is roughly constant along the length of these bands and varies systematically from one band to the next. A confusingly large number of different response characteristics are thought to be ‘mapped’ along the iso-frequency bands [3,4]. These include response threshold, the dynamic range and shape of response-level functions, the bandwidth and shape of frequency-response profiles, sensitivity to frequency modulation, and the type of binaural interaction exhibited by the neurons. The diverse and variable nature of these superimposed representations makes the task of finding the right stimulus to study any particular neuron a daunting one.

Finding the right stimulus

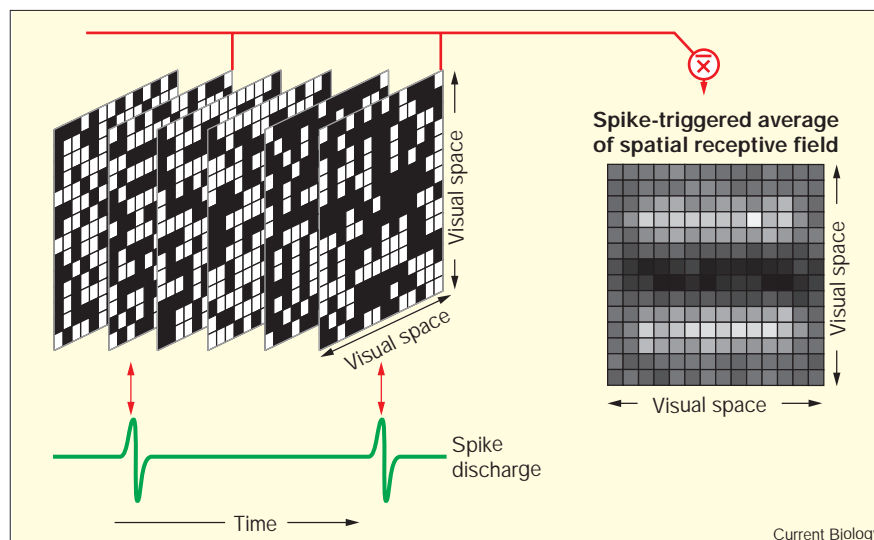
As studies of the forebrain song-system in birds [5] and of the auditory cortex in echolocating bats [6] have demonstrated, considerable progress can be made if the choice of stimuli is based on behavioural considerations. But it is a lot harder to make an ‘inspired guess’ as to what the appropriate stimuli might be in less specialised sensory systems. Indeed, the discovery of orientation selectivity in V1 was made quite by accident [7].

Some studies have used search algorithms in which a computer continually changes the stimulus on the basis of the neuron’s response [8]. But this approach is often hampered by the variable, stochastic nature of neural responses and by the ‘high dimensionality of the search space’ — that is, the large number of different ways in which the stimulus could be altered in an attempt to make it more effective. An alternative strategy is therefore to use very large numbers of stimuli and a robust statistical technique to deal with the stochastic nature of the responses. One technique that does this is known as ‘reverse-correlation’, and it has recently been applied successfully to the study of A1 neurons [9].

In sensory physiology experiments, stimuli are typically presented in isolation and the experimenter looks for a relationship between stimulus parameter and neural

Figure 1

The 'reverse-correlation' approach to determining the optimum visual stimulus for a neuron in the visual cortex. A sequence of many randomly flickering checkerboard patterns is presented in quick succession. The neuron will respond to flickering squares that fall within its spatial receptive field, as indicated by the spike discharges in the lower part of the figure. Note that only two of the patterns shown are effective in exciting the neuron sufficiently to generate a spike discharge. By averaging the stimuli that precede these spikes, one can build up an estimate of the spatial receptive field and therefore of the optimal stimulus for the neuron. Many different checkerboard patterns are needed to do this; in this simulation, 10,000 randomized patterns were used. The spatial receptive field of the neuron contains a horizontally-oriented inhibitory region (dark squares) that is flanked on either side by an excitatory region (light squares). The reverse-correlation technique therefore predicts that the optimal stimulus for this simple cell in V1 is a horizontal dark bar that falls within the



inhibitory region. By averaging the stimuli at different intervals preceding the spikes to generate a spatio-temporal receptive field (not

shown), it is possible to show how the selectivity of the neuron changes with time.

response. In reverse-correlation experiments, however, stimuli are presented in a continuous stream of random patterns or sounds, and post hoc attempts are made to correlate responses with stimulus features. The first reverse-correlation experiments [10] studied auditory nerve responses using continuous white noise, which contains sound energy at all frequencies but randomly fluctuating energy levels. Reverse-correlation experiments on the visual system often use sequences of random patterns of bright and dark patches to generate 'visual noise' similar to that seen on a de-tuned TV-set (Figure 1). Most of the random patterns presented will be ineffective, but some trigger the firing of action potentials or 'spikes'. The investigator then tries to identify what all the random stimulus episodes preceding the neuron's discharges had in common, typically by calculating the 'spike triggered average' of all the stimulus patterns preceding a discharge.

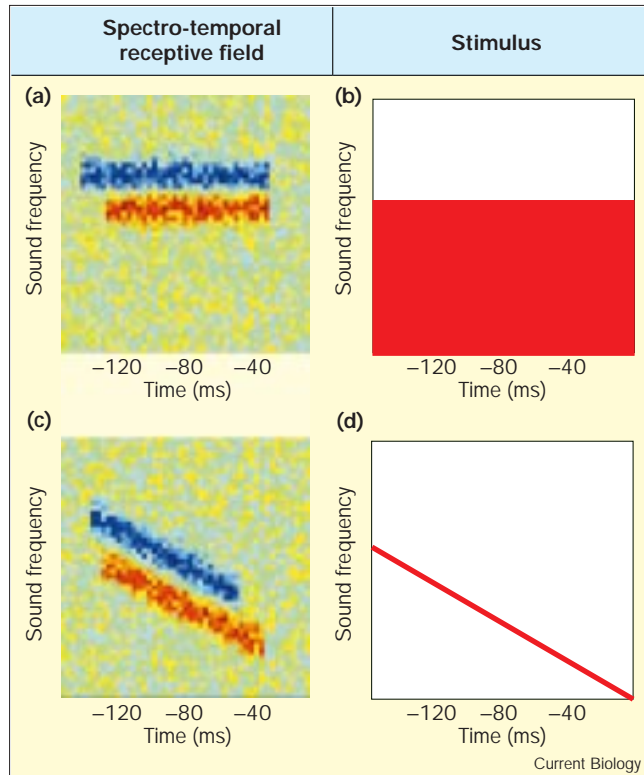
Reverse-correlation techniques have some advantages over more traditional 'forward-correlation' experiments. Two of these are beautifully illustrated in the data obtained by deCharms *et al.* [9] in their recent study of A1. The first advantage is that, as stimuli are presented in a continuous stream in reverse-correlation experiments, it is sometimes possible to reveal temporal structure in a neuron's receptive field. To do this, spike-triggered averages are calculated, not just for the stimulus events at one particular time delay prior to the neuron's discharges, but for a series of intervals extending back in time from the moment of discharge. For auditory neurons, the result

of this analysis is referred to as the 'spectro-temporal receptive field'.

The second advantage of the reverse-correlation approach is that it can be quite effective at revealing inhibitory areas in a neuron's receptive field. Because A1 neurons typically exhibit little spontaneous activity, it is difficult to identify inhibitory regions by presenting isolated tones. In contrast, the spectro-temporal receptive fields that deCharms *et al.* [9] calculated from neural responses to random tone combinations often revealed tones in particular frequency bands whose absence was associated with individual spikes. In other words, the presence of such tones reduced neuronal firing, indicating that these frequency bands were inhibitory. Some of the spectro-temporal receptive fields thus obtained displayed alternating excitatory and inhibitory bands (Figure 2), not unlike those seen in the receptive fields of simple cells in V1 (Figure 1).

Neural filters and feature extractors

Because reverse-correlation uses very large numbers of stimuli chosen at random, this approach is less constrained than other methods by prior assumptions about the nature of the optimal stimulus. There are, however, several limitations inherent in the reverse-correlation technique. Perhaps the most serious of these is that the neurons must, at least very approximately, behave like 'linear filters' for the stimulus parameters under investigation. If a hypothetical neuron responds best to a particular pattern, say the letter 'B', and is linear, then it would respond half maximally when presented with half a letter

Figure 2

(a,c) Spectro-temporal receptive fields of two neurons in the primary auditory cortex of a primate, estimated using the reverse-correlation technique. Neuronal activity is measured during stimulation with a sequence of rapidly changing chords. The spectro-temporal receptive field is constructed by averaging the spectral content of the stimulus episodes preceding each action potential recorded. The colour code indicates the effectiveness of different frequency bands as a function of time preceding the occurrence of the action potentials. Warm colours (red and yellow) indicate excitatory regions of the spectro-temporal receptive field, and cool colours (blue) indicate inhibitory regions. The spectro-temporal receptive field shown in (a) exhibits a narrow, constant-frequency region of excitation, whereas that in (c) shows a single excitatory region that shifts in frequency with time. In both cases, a higher-frequency region of inhibition flanks the excitatory region. The spectro-temporal receptive field in (a) predicts that this neuron will respond preferentially to narrow frequency bands or to constant-frequency ‘edges’ that correspond to the excitatory region. Low-pass noise (red) with an upper cut-off at this frequency (b) is a highly effective stimulus, leading to the notion that A1 neurons of this type are spectral edge detectors [9]. Similarly, the neuron illustrated in (c) responds selectively to tonal stimuli that are swept in frequency in direction and at a rate (d) that match the excitatory region of the spectro-temporal receptive field.

B. This property is essential for the reverse-correlation technique to work, as we rely on the neuron’s discharges to pick out all those stimuli from our random stimulus set that happen to contain fragments of the letter B in order to piece together the complete pattern in the spike-triggered average. Half a letter B may, of course, be mistaken for either an ‘E’ or a ‘D’. A genuine ‘B-detector’ should be capable of distinguishing these letters categorically, but it

would either not respond to random dot patterns or do so in a highly non-linear manner. A reverse-correlation analysis of the neuron’s response based on such stimuli would therefore almost certainly fail to reveal its optimal pattern.

If a neuron were to act like a linear filter, however, then its spike-triggered average would converge to an exact description of the optimal stimulus. As deCharms *et al.* [9] were able to construct auditory stimuli on the basis of spectro-temporal receptive fields that evoked considerably stronger responses than those usually observed in A1, it seems likely that at least some A1 neurons — like many in V1 [11] — behave approximately like linear filters for particular stimulus attributes. To illustrate the analogy between A1 and V1, deCharms *et al.* [9] showed that A1 neurons with spectro-temporal receptive fields that contain a narrow, constant-frequency excitatory region bordered by inhibition respond well to ‘low-pass noise’ — a stimulus containing sound energy only at frequencies below a certain value — with a cut-off frequency that matches the excitatory region (Figure 2a,b), but much less well to stimuli with other cut-off frequencies. This is comparable to the effect of placing an appropriately oriented edge over the receptive field of a V1 neuron.

deCharms *et al.* [9] also described spectro-temporal receptive fields in which excitatory and inhibitory regions vary in frequency with time, indicating that the neurons are selective for the rate and direction of frequency modulation (Figure 2c,d). They suggested that the auditory neurons with spectro-temporal receptive fields of this type might be equivalent to visual neurons that exhibit a preferred direction of motion. The A1 neurons characterized in this study are referred to as detectors of stimulus ‘edges’ in either frequency or time. This terminology, however, implies that they are feature extractors. In fact, it is probably inappropriate to describe neurons at this early stage of cortical processing — whether they are found in A1 or V1 — as edge detectors, rather than just simple linear filters.

While these findings imply that comparable linear processing strategies may be employed in V1 and A1, it does not follow that the non-linear response characteristics of other neurons in the two systems will be the same. Moreover, we must not forget that there are many profound differences between the structures of the visual and the auditory worlds. For example, visual objects are bounded by edges and edge detection can be very useful for visual scene segmentation. Whereas sound onsets — ‘temporal edges’ — play a major role in auditory scene analysis [12], the contribution of spectral edges in sounds is much less clear [13].

Beyond single neurons

Nevertheless, the analogies drawn by deCharms *et al.* [9] between the stimulus processing performed by neurons in

A1 and by those in V1 raise a number of important questions. For example, are A1 neurons that exhibit a particular form of spectro-temporal receptive field arranged into discrete columns of the sort found throughout V1? Do these response characteristics vary in a systematic manner over the surface of the cortex? Do they emerge within the cortex as a result of converging input patterns or through intracortical circuitry? Issues like these have provided the focus for extensive research on the visual cortex for the past 40 years, and it seems likely that the study by deCharms *et al.* [9] will become a catalyst for future auditory experiments.

Another major feature of the visual cortex is that functionally specialized processing streams related to different classes of retinal-ganglion-cell input emerge from distinct compartments within V1 [14]. After varying degrees of crosstalk, these streams project either ventrally to the inferotemporal cortex or dorsally to the posterior parietal cortex, where they appear to mediate object recognition and visuomotor control, respectively [15]. On the basis of the rather imprecise spatial segregation of response properties along the iso-frequency dimension of A1, it has been suggested that different aspects of auditory processing — frequency analysis, intensity discrimination, sound localization and so on — may take place in parallel [4]. If this follows the same principle as in V1, then we should find that functionally-distinct channels convey selective information to the cortical fields that lie beyond A1.

The notion that non-primary cortical fields have distinct roles in processing biologically important acoustical signals certainly receives ample support from studies of the echolocating mustached bat [6]. Whether this is also the case in less specialized mammals is largely unknown [16], although tools, such as reverse-correlation, that help to identify the optimal stimulus for auditory neurons may be useful in this endeavour.

Acknowledgements

We are grateful to Bruce Cumming, Ian Thompson and Roland Baddeley for their helpful comments on this paper. Andrew King is a Wellcome Senior Research Fellow and Jan Schnupp is a Dunhill Medical Trust Research Fellow.

References

1. Hubel DH, Wiesel TN: **Receptive fields, binocular interaction and functional architecture in the cat's visual cortex.** *J Physiol Lond* 1962, **160**:106-154.
2. Hübener M, Shoham D, Grinvald A, Bonhoeffer T: **Spatial relationships among three columnar systems in cat area 17.** *J Neurosci* 1997, **17**:9270-9284.
3. Ehret G: **The auditory cortex.** *J Comp Physiol A* 1997, **181**:547-557.
4. Schreiner CE: **Spatial distribution of responses to simple and complex sounds in the primary auditory cortex.** *Audiol Neurootol* 1998, **3**:104-122.
5. Doupe AJ, Konishi M: **Song-selective auditory circuits in the vocal control system of the zebra finch.** *Proc Natl Acad Sci USA* 1991, **88**:11339-11343.
6. Suga N, Yan J, Zhang Y: **Cortical maps for hearing and egocentric selection for self-organization.** *Trends Cog Sci* 1997, **1**:13-20.

7. Hubel DH: *Eye, Brain, and Vision.* New York: Scientific American Library; 1988.
8. Nelken I, Prut Y, Vaadia E, Abeles M: **In search of the best stimulus: an optimization procedure for finding efficient stimuli in the cat auditory cortex.** *Hear Res* 1994, **72**:237-253.
9. deCharms RC, Blake DT, Merzenich MM: **Optimizing sound features for cortical neurons.** *Science* 1998, **280**:1439-1443.
10. de Boer E, de Jongh HR: **On cochlear encoding: potentialities and limitations of the reverse-correlation technique.** *J Acoust Soc Am* 1978, **63**:115-135.
11. Movshon JA, Thompson ID, Tolhurst DJ: **Spatial summation in the receptive fields of simple cells in the cat's striate cortex.** *J Physiol Lond* 1978, **283**:53-77.
12. Mellinger DK, Mont-Reynaud BM: **Scene analysis.** In *Auditory Computation.* Edited by Hawkins HL, McMullen TA, Popper AN, Fay RR. New York: Springer-Verlag; 1996:271-331.
13. Small AM, Daniloff RG: **Pitch of noise bands.** *J Acoust Soc Am* 1967, **41**:506-512.
14. Ts'o DY, Roe AW: **Functional compartments in visual cortex: segregation and interaction.** In *The Cognitive Neurosciences.* Edited by Gazzaniga MS. Cambridge, Massachusetts: MIT Press; 1995:325-337.
15. Goodale MA: **Visual pathways supporting perception and action in the primate cerebral cortex.** *Curr Opin Neurobiol* 1993, **3**:578-585.
16. Rauschecker JP, Tian B, Pons T, Mishkin M: **Serial and parallel processing in rhesus monkey auditory cortex.** *J Comp Neurol* 1997, **382**:89-103.